CS 134: Networks	Lecture 4, Wednesday, 2/1
Prof. Yaron Singer	Spring 2017

1 Overview

In the last lecture we discussed Milgram's experiment and the small-world phenomenon. In search of mathematical models of social networks that can describe this phenomenon, we emphasized three important properties.

Desirable properties of social network models:

1. The *average degree* of the nodes in the graph is **constant**;
2. The *clustering coefficient* of the graph is **large** (constant, bounded from 0);
3. The *diameter* of the graph is **small** (poly-logarithmic in size of graph).

We constructed a model of small-world networks called the Watts-Strogatz model and showed that it satisfies all the above properties. While the model guarantees that short paths in the network exist, it still does not fully explain Milgram's experiment. Milgram's experiment shows that not only do short paths exist, but that people were very good at finding short paths, using very little information about the network. While participants presumably knew who their acquaintances were and where they lived, they did not the entire social network. How were these participants able to find a path to Cambridge without knowledge of the social network structure?

2 Navigation in a Small World

To address the question of how individuals were able to *navigate* messages with only local information, Kleinberg [1] suggests a small-world model defined as follows.

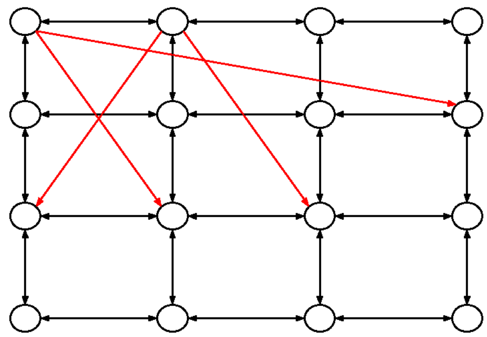

Figure 1: A partially constructed Kleinberg model with $k = 1$, $\ell = 2$, and $n = 16$. We start with a 4×4 square lattice. Connect all nodes with a lattice distance of 1 (black arrows). Then for each node u, add 2 random edges where the probability of adding edge (u, v) is proportional to $\Delta(u, v)^{-2}$. For clarity, the diagram above shows only a partial construction – only the first two nodes on the first row have been augmented with random edges (red arrows). Note also that Kleinberg's model is a random graph, so on a different realization the black arrows will stay the same but the red arrows will likely change.

Definition. *The Kleinberg model is defined as follows:*

- *Nodes are positioned on a square lattice with n nodes.*
- *Each node is connected to the nodes that are at most a distance k away on the lattice, for some given parameter $k \in \mathbb{N}$.*
- *For some given $\ell \in \mathbb{N}$ each node u has ℓ* long range *edges which are randomly connected to other nodes in the network with the following probability:*

$$\Pr[u \to v] = \frac{\Delta(u, v)^{-2}}{\sum_{w \neq u} \Delta(u, w)^{-2}}$$

where

  - *$u \to v$ is the event of a long range edge from u to v*
  - *$\Delta(u, v)$ is the distance from u to v on the grid*

2.1 Properties of Kleinberg's model

Lemma 1. *In Kleinberg's model,*

1. *for all $u \in V$, the number of nodes at lattice distance exactly d from u is $\Theta(d)$;*
2. *for all $u \in V$, the number of nodes at lattice distance at most d from u is $\Theta(d^2)$;*
3. $\Pr[u \to v] = \Theta(\frac{\Delta(u,v)^{-2}}{\log n})$.

Proof.

1. For the first property, consider a node v. If v's coordinates are $(i, j) \in n \times n$ and w.l.o.g. $i, j > d$, then the d nodes $(i-d, j), (i-(d-1), j+1), \ldots, (i, j-d)$ are all at lattice distance exactly d. And this happens at most in four directions on the grid.

2. Given the first property we see that $\sum_{i=1}^{d} i = \frac{d(d+1)}{2} = \Theta(d^2)$ which gives us the second property.

3. Finally, the third property $\Pr[u \to v] = \frac{\Delta(u,v)^{-2}}{\sum_{w \neq u} \Delta(u,w)^{-2}}$ is due to the fact that in the denominator, by partitioning the sum based on the lattice distance of the nodes from u,

$$\sum_{w \neq u} \Delta(u, w)^{-2} = \Theta\left(\sum_{i=1}^{2(\sqrt{n}-1)} i \cdot i^{-2}\right) = \Theta\left(\sum_{i=1}^{2(\sqrt{n}-1)} i^{-1}\right) = \Theta(\log n^{1/2}) = \Theta(\log n).$$

$\square$

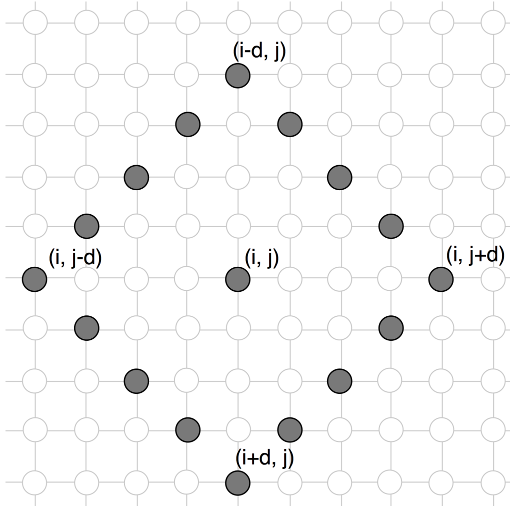

Figure 2: A square lattice near coordinate (i, j). The outer shaded vertices have a lattice distance exactly d from (i, j). The vertices inside this square have a lattice distance less than d from $(i.j)$.

3 The Main Result

Given a message held by an individual that seeks to forward it to some target destination, a reasonable heuristic to use is simply to forward the message to the person closest to the target. Remarkably, this simple and intuitive heuristic explains the shortest paths observed in Milgram's experiment.

Theorem (Main Result)**.** *For graphs generated according to Kleinberg's model, even with $k = \ell = 1$, when each node forwards the message to its neighbor that is closest to the target, the message reaches its destination after $O(\log^2 n)$ steps, in expectation.*

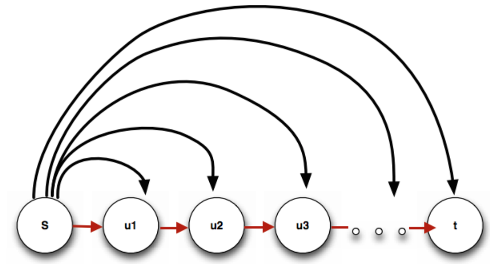

Figure 3: A graph for random routing to act on. We wish to route from s to t, red edges are fixed and black edges are random.

A Note on routing on random graphs. Instead of generating a random graph and analyzing the algorithm, we can generate the long-range links from a node u only once the algorithm visits node u since no node is visited more than once. Using the principle of deferred decisions, we have that analyzing the deterministic routing on the random graph is equivalent to analyzing a randomized algorithm on a fixed graph. For a more detailed discussion of this concept, please refer to the corresponding section notes.

3.1 Proving the theorem

We can now discuss the proof of the theorem. The message passing rule is the following:

message passing rule: pass the message to the neighbor who is closest to the target.

Remarkably, this natural and simple rule enables the main result.

Proof idea. The main idea behind the proof is the following. At any given step, after $\log n$ further steps (in expectation), the remaining distance to the target is halved. It will keep doing this until it reaches the destination. Since the distance is at most $O(n)$ between a source and a destination, after halving the distance $O(\log n)$ times, the message will reach its destination. This is where the $\log^2 n$ comes from. We'll prove this next.

Proof of the main theorem, We say that the Algorithm is at phase j if the message is held by u s.t. $2^j < \Delta(u,t) \le 2^{j+1}$. We will let X_j be the random variable denoting the number of steps at phase j.

Lemma 2. $\mathbf{E}[X_j] = O(\log n)$.

Proof. Let u be a node on the lattice and suppose the algorithm is at phase j. First, note that for $0 \le j \le \log(\log n)$ the algorithm will spend at most $\log n$ steps in phase j since the next phase is at distance at most $2^{\log(\log n)} = \log n$ on the grid.

So we will now consider the case in which $j \ge \log(\log n)$. Define ball $B_j = \{v | \Delta(v,t) \le 2^j\}$ to be all the nodes that are in one of the next phases $i < j$. Given a current node u outside B_j, we want to know what is the probability that u has a long range edge to a node inside B_j, which we denote by $u \to B_j$. We use v^* to denote the node that is the furthest away from u in B_j, that is

$v^* := \arg\max_{v \in B_j} \Delta(u, v)$. Then,

$$\begin{aligned}
\Pr(u \to B_j) &= \sum_{v \in B_j} \Pr(u \to v) \\
&\geq \sum_{v \in B_j} \Pr(u \to v^*) \\
&= \Pr(u \to v^*) \, |B_j| \\
&= \Theta\left(\frac{\Delta(u, v^*)^{-2}}{\log n} |B_j| \right) \\
&= \Theta\left(\frac{\Delta(u, v^*)^{-2} 2^{2j}}{\log n} \right)
\end{aligned}$$

where the inequality is since each vertex in B_j is at least as likely to be jumped to as v^*, the third equality is due to the third property from Lemma 1, and the last equality is due to the second property from Lemma 1. Now observe that

$$\begin{aligned}
\frac{\Delta(u, v^*)^{-2} 2^{2j}}{\log n} &\geq \frac{(\Delta(u, t) + \Delta(t, v^*))^{-2} 2^{2j}}{\log n} \\
&\geq \frac{(2^{j+1} + 2^j)^{-2} 2^{2j}}{\log n} \\
&\geq \frac{(3 \cdot 2^j)^{-2} 2^{2j}}{\log n} \\
&= \Theta\left(\frac{1}{\log n} \right)
\end{aligned}$$

where the first inequality is by the triangle inequality and the second inequality is by by definition of phase j and by definition of ball B_j. Therefore, $\Pr(u \to B_j) = \Omega\left(\frac{1}{\log n}\right)$. Finally, the expected number of steps at phase j is

$$\begin{aligned}
\mathbf{E}[X_j] &= \sum_{i=1}^{\infty} i \cdot \Pr(X_j = i) \\
&= \sum_{i=1}^{\infty} \Pr(X_j \geq i) \\
&= \sum_{i=1}^{\infty} \left(1 - \Omega\left(\frac{1}{\log n} \right) \right)^{i-1} \\
&= O(\log n)
\end{aligned}$$

since the probability that phase j lasts longer than i steps is equal to the probability that at each of the $i - 1$ first steps, there are no long range edges in B_j and since $\sum_{i=1}^{\infty} x^i = \frac{x}{1-x}$ when $-1 < x < 1$. $\square$

To conclude the proof of the main result, let X be the total number of steps and observe that there are $\log n$ stages. Then,

$$\begin{aligned}
\mathbf{E}[X] &= \mathbf{E}\left[\sum_{j=1}^{\log n} X_j\right] \\
&= \sum_{j=1}^{\log n} \mathbf{E}[X_j] \\
&= \sum_{j=1}^{\log n} O(\log n) \\
&= O(\log^2 n).
\end{aligned}$$

References

[1] Jon Kleinberg. *The small-world phenomenon: An algorithmic perspective.* Proc. 32nd ACM Symposium on Theory of Computing, 2000.